

# On the existence of compact $\varepsilon$ -approximated formulations for knapsack in the original space

Yuri Faenza\*

Laura Sanità†

## Abstract

We show that there exists a family  $\mathcal{P}$  of Knapsack polytopes such that for each  $P \in \mathcal{P}$  and each  $\varepsilon > 0$ , any  $\varepsilon$ -approximated formulation of  $P$  in the original space  $\mathbb{R}^n$  requires a number of inequalities that is super-polynomial in  $n$ . This answers a question by Bienstock and McClosky (2012). We also prove that, for any down-monotone polytope, an  $\varepsilon$ -approximated formulation in the original space can be obtained with inequalities using at most  $O(\frac{1}{\varepsilon} \min\{\log(n/\varepsilon), n\})$  different coefficients.

## 1 Introduction

The *Knapsack* problem is one of the classics of combinatorial optimization: given a set of  $n$  objects, each coming with a nonnegative weight and profit, and a threshold  $B$ , find the most profitable subset of objects whose total weight does not exceed  $B$ . Knapsack has been extensively studied in the combinatorial optimization and integer programming communities (see e.g. [9, 12]).

From an algorithmic point of view, we know that the associated decision problem is NP-Complete [8]. On the sunny side, the Knapsack problem admits a *fully polynomial-time approximation scheme* (FPTAS) [6, 11]. That is, for each  $\varepsilon \in \mathbb{R}_+$  and  $n \in \mathbb{N}$ , a feasible solution to a Knapsack instance on  $n$  objects of cost at least  $(1 - \varepsilon)$  the cost of the optimum can be found in time polynomial in  $1/\varepsilon$  and  $n$ . Roughly speaking, this implies that we can find an approximated solution with cost arbitrarily close to the one of the optimum, in time polynomial in the input size and in the accuracy of the approximation. This algorithm relies on classical combinatorial techniques such as scaling, rounding, and dynamic programming.

A natural question [2, 3, 15] is whether we can translate the *algorithmic* approximation given by the FPTAS into a *polyhedral* approximation, that is, whether we can approximate the *Knapsack polytope* using a polynomial number of inequalities, where the Knapsack polytope is the convex hull of all 0 – 1 vectors in  $\mathbb{R}^n$  corresponding to feasible solutions of a given Knapsack instance. Let us formalize the concept of polyhedral approximation by introducing some definitions.

**Definition 1.** *Given a polytope  $P \subseteq \mathbb{R}^n$  and  $\varepsilon > 0$ , we say that  $Q \subseteq \mathbb{R}^n$  is an  $\varepsilon$ -approximated formulation for  $P$  if  $P \subseteq Q$ , and for each  $c \in \mathbb{R}^n$  the following inequality holds*

$$\max\{cx : x \in P\} \geq (1 - \varepsilon) \max\{cx : x \in Q\}. \quad (1)$$

---

\*DISOPT, Ecole Polytechnique Fédérale de Lausanne (Switzerland): [yuri.fienza@epfl.ch](mailto:yuri.fienza@epfl.ch). Supported by the *Ambizione* grant PZ00P2\_154779 *Tight formulations of 0-1 problems* funded by the Swiss National Science Foundation.

†Combinatorics and Optimization department, University of Waterloo (Canada): [lsanita@uwaterloo.ca](mailto:lsanita@uwaterloo.ca)

For a polytope  $Q \subseteq \mathbb{R}^n$ , we denote by  $\text{size}(Q)$  the number of inequalities in a minimum linear description of  $Q$  (i.e. the number of *facets* of  $Q$ ). We also denote by  $xc(Q)$  its *extension complexity*, that is the minimum size of a polytope  $\tilde{Q}$  such that there exists a linear map  $\pi$  with  $Q = \pi(\tilde{Q})$ . Note that  $Q$  and  $\tilde{Q}$  can have different dimensions.

A family  $\mathcal{P}$  of polytopes is said to have a *polynomial-size relaxation schemes* (PSRS) if for each  $\varepsilon > 0$  there exists a polynomial function  $\phi : \mathbb{N} \rightarrow \mathbb{N}$  such that for each  $P \in \mathcal{P}$  with  $P \subseteq \mathbb{R}^n$ ,  $P$  has an  $\varepsilon$ -approximated formulation  $Q$  with  $xc(Q) \leq \phi(n)$ . If, in addition, the function  $\phi$  grows polynomially in  $\frac{1}{\varepsilon}$ , then  $\mathcal{P}$  is said to have a *fully polynomial-size relaxation schemes* (FPSRS).

The above definitions (see [4]) of PSRS and FPSRS can be seen as a translation of the algorithmic notions of FPTAS and PTAS into polyhedral terms.

For the family of Knapsack polytopes, the existence of a PSRS has been proved by Bienstock [2], while the existence of a FPSRS is still an interesting open question [15, 2].

The PSRS provided by [2] yields an  $\varepsilon$ -approximated formulation  $Q$  with small extension complexity, but it does not imply strong bounds on  $\text{size}(Q)$ . In fact, in the same paper Bienstock [2], as well as Bienstock and McClosky [3, Question 1, page 340], ask whether such result can be straightened by proving the existence of a PSRS *in the original space*, that is given by replacing  $xc(Q)$  with  $\text{size}(Q)$  in the definition of PSRS. In other words, whether for any Knapsack polytope  $P \subseteq \mathbb{R}^n$  we can find an  $\varepsilon$ -approximated formulation  $Q \subseteq \mathbb{R}^n$  with a number of *facets* bounded by a polynomial function  $\phi(n)$ .

In this paper, we negatively answer this open question.

**Theorem 1.** *The family of Knapsack polytopes admits no PSRS in the original space.*

Bienstock and McClosky [3] also ask for upper bounds on the number of inequalities needed to obtain an  $\varepsilon$ -approximation in the original space for the Knapsack polytope. To the best of our knowledge, the only known upper bound is the one of Van Vyve [14], cited in [3], of order  $(\lceil n/\varepsilon \rceil + 1)^n$ . This follows from the fact that the formulation containing all valid inequalities with integer coefficients between 0 and  $\lceil n/\varepsilon \rceil$  is  $\varepsilon$ -approximated. In [3], it is also shown that, for each  $\gamma > 0$ , there exist Knapsack polytopes  $P$  such that the polytope  $Q$  given by all valid inequalities with integer coefficients up to  $n^{1-\gamma}$  is a bad approximation. Namely, there exists an objective function  $c \in \mathbb{R}_+$  such that the ratio (1) is arbitrarily close to  $1/2$ .

Our second result is a new upper bound on the number of different coefficients (hence on the number of facets) needed to obtain an  $\varepsilon$ -approximated formulation for Knapsack polytopes in the original space. Our upper bound improves over the bound in [14] whenever  $1/\varepsilon$  is a sub-exponential function of  $n$ , and asymptotically matches it otherwise. Moreover, our result holds for any *down-monotone* polytope, that is, any polytope  $P \subseteq \mathbb{R}_+^n$  with the property that  $x \in P$  and  $0 \leq y \leq x$ , implies  $y \in P$ .

**Theorem 2.** *For any down-monotone polytope  $P \subseteq \mathbb{R}_+^n$ , an  $\varepsilon$ -approximated formulation in the original space can be obtained with inequalities using at most  $O(\frac{1}{\varepsilon} \min\{\log(n/\varepsilon), n\})$  different coefficients. Each of those coefficients is an integer between 0 and  $5n(1 - \varepsilon)/\varepsilon$ . In particular, there exists an  $\varepsilon$ -approximated formulation with  $O(\frac{1}{\varepsilon} \min\{\log(n/\varepsilon), n\})^n$  facets.*

We emphasize that Theorem 2 implies the existence of  $\varepsilon$ -approximated formulations of a down-monotone polytope with only  $O(\log n)$  different coefficients, if  $\varepsilon$  is *fixed*.

## 2 No PSRS in the original space exists for Knapsack

The goal of this section is to prove Theorem 1.

**The strategy.** For any given  $\varepsilon > 0$  small enough, we provide an infinite sequence of integers, and to each integer  $n$  from this sequence we associate a Knapsack polytope  $P(n, \varepsilon) \subseteq \mathbb{R}^n$  and a family of points  $\mathcal{X}(n, \varepsilon) \subseteq \mathbb{R}^n$  with three properties:

- (i)  $|\mathcal{X}(n, \varepsilon)|$  is super-polynomial in  $n$ ;
- (ii) for each  $\bar{x} \in \mathcal{X}(n, \varepsilon)$ , there exists  $c \in \mathbb{R}_+^n$  such that  $(1 - \varepsilon)c\bar{x} > \max\{cx : x \in P(n, \varepsilon)\}$ ;
- (iii) for each pair of distinct points  $\bar{x}_1, \bar{x}_2 \in \mathcal{X}(n, \varepsilon)$ , the line segment between  $\bar{x}_1$  and  $\bar{x}_2$  intersects  $P(n, \varepsilon)$ .

In particular, (ii) implies that  $P(n, \varepsilon) \cap \mathcal{X}(n, \varepsilon) = \emptyset$  and, most important, that any  $\varepsilon$ -approximation of  $P(n, \varepsilon)$  must have an inequality separating each  $\bar{x} \in \mathcal{X}(n, \varepsilon)$  from  $P(n, \varepsilon)$ . However, (iii) implies that no two points  $\bar{x}_1, \bar{x}_2 \in \mathcal{X}(n, \varepsilon)$  can be separated by *the same* inequality, since the line segment between them intersects  $P(n, \varepsilon)$ . It follows that any description of an  $\varepsilon$ -approximation of  $P(n, \varepsilon)$  contains at least  $|\mathcal{X}(n, \varepsilon)|$  distinct inequalities, that is super-polynomial in  $n$  by (i). As  $\varepsilon$  was arbitrary and the choices of  $n$  were infinite, this implies Theorem 1. We remark that similar strategies have been recently used by Kaibel and Weltge [7] and Kolliopoulos and Moysoglou [10].

**Constructing a family of Knapsack polytopes.** Let  $\varepsilon > 0$  be a fixed small number  $< \frac{2}{27}$ . For any  $n > 1/\varepsilon$ ,  $n \geq 130$  that is the square of a prime number, we construct a Knapsack polytope as follows:

$$P(n+1, \varepsilon) = \{x \in \{0, 1\}^{n+1} : \sum_{i=1}^n \frac{1}{2\varepsilon\sqrt{n}}x_i + nx_{n+1} \leq n + \frac{1}{2\varepsilon} - 1\},$$

Interestingly, note that  $P(n+1, \varepsilon)$  describes a somewhat “easy” Knapsack instance. In fact, the set of feasible solutions can be partitioned into two subsets: those with  $x_{n+1} = 1$  and those with  $x_{n+1} = 0$ . If we fix the value of  $x_{n+1}$ , the residual problem in both cases reduces to a Knapsack instance where all the  $n$  remaining objects have the same size. Therefore a greedy algorithm can be used to find an optimal solution to each subproblem. Similarly, from a polyhedral point of view, one easily realizes that the convex hull of the points in  $P(n+1, \varepsilon)$  admits a compact extended formulation. This can be obtained as follows: partition again the family of feasible solutions into two sets based on the value of  $x_{n+1}$ . As one immediately checks, the convex hull of each of those two subproblems has a compact formulation. Therefore one can apply Balas’ union of polyhedra technique [1] to obtain the claimed compact extended formulation. However, the situation drastically change when we require our formulation to yield a polytope in  $\mathbb{R}^{n+1}$ . In this case, we can not even  $\varepsilon$ -approximate our polytope without using a super-polynomial number of inequalities, as we are now going to show.

**Constructing a family of points.** Our goal is to define a family of points  $\mathcal{X}(n+1, \varepsilon)$  that together with  $P(n+1, \varepsilon)$  satisfy properties (i)-(iii). The main idea lies in realizing that this can be achieved by constructing a suitable *set system*  $\mathcal{S}$  of subsets of  $[n]$ , containing a super-polynomial number of sets with the property that the sets have pairwise small intersection.

A way to construct such a set system is following an idea of Nisan and Wigderson [13, Lemma 2.5]. Consider the field  $\mathbb{F}$  with  $\sqrt{n}$  elements, which exists because we are assuming  $\sqrt{n}$  to be a prime number. Consider the universe set  $U$  made of  $n$  elements represented as ordered pairs  $(a, b)$  with  $a, b \in \mathbb{F}$ . To each polynomial  $\pi$  of  $\mathbb{F}$  of degree at most  $\lfloor \sqrt{n}/2 - 4 \rfloor$ , we associate the set  $S_\pi = \{(a, \pi(a))\}_{a \in \mathbb{F}}$ . Note that  $|S_\pi| = \sqrt{n}$ . Let  $\mathcal{S}$  be the family of all such sets. It is well-known (see e.g. [5]) that, for each choice of  $2(k+1)$  elements  $a_1, \dots, a_{k+1}, b_1, \dots, b_{k+1} \in \mathbb{F}$ , with the first  $k+1$  being distinct, there exists exactly one polynomial  $\pi$  in  $\mathbb{F}$  of degree at most  $k$  such that  $\pi(a_i) = b_i$  for each  $i$ . We then deduce  $|\mathcal{S}| = n^{\frac{1}{2}(\lfloor \sqrt{n}/2 - 4 \rfloor + 1)}$  and  $|S_\pi \cap S_{\pi'}| \leq$

$\lfloor \sqrt{n}/2 - 4 \rfloor$  for each pair of polynomials  $\pi \neq \pi'$  of  $\mathbb{F}$ . Hence, we obtained a family  $\mathcal{S} \subseteq 2^n$  such that:

- (a)  $|S| = \sqrt{n}$  for each  $S \in \mathcal{S}$ ,
- (b)  $|S \cap S'| \leq \lfloor \sqrt{n}/2 - 4 \rfloor$  for each  $S \neq S' \in \mathcal{S}$ ,
- (c)  $|\mathcal{S}| = n^{\frac{1}{2}(\lfloor \sqrt{n}/2 - 4 \rfloor + 1)}$ .

From now on, we denote the elements of  $U$  with numbers 1 to  $n$ , and identify the  $i$ -th element with the  $i$ -th object from the knapsack instance, for  $i = 1, \dots, n$ . For each  $S \in \mathcal{S}$  we define a point  $x^S \in \mathbb{R}^{n+1}$  with coordinates:

- $x_{n+1}^S = \frac{4}{5}$ ,
- $x_i^S = 1 - \frac{\varepsilon}{2}$  for  $i \in S$ ,
- $x_i^S = 0$  for  $i \notin S$  and  $i \neq n+1$ .

We let  $\mathcal{X}(n+1, \varepsilon)$  be the family of these points, that is  $\mathcal{X}(n+1, \varepsilon) := \{x^S : S \in \mathcal{S}\}$ .

**Verifying properties.** We now show that  $\mathcal{X}(n+1, \varepsilon)$  and  $P(n+1, \varepsilon)$  satisfy properties (i), (ii) and (iii), concluding our proof.

**Claim 1.**  $\mathcal{X}(n+1, \varepsilon)$  satisfies (i).

*Proof.*  $|\mathcal{X}(n+1, \varepsilon)| = |\mathcal{S}|$  and, by (c),  $|\mathcal{S}| = n^{\frac{1}{2}(\lfloor \sqrt{n}/2 - 4 \rfloor + 1)}$ . □

**Claim 2.**  $\mathcal{X}(n+1, \varepsilon)$  and  $P(n+1, \varepsilon)$  satisfy (ii).

*Proof.* Fix  $S \in \mathcal{S}$ , and define  $c \in \mathbb{R}^n$  as follows:

- $c_{n+1} = 1$ ,
- $c_i = \frac{1}{2\varepsilon\sqrt{n}}$  for  $i \in S$ ,
- $c_i = 0$  for  $i \notin S$  and  $i \neq n+1$ .

We first show that

$$\max\{cx : x \in P(n+1, \varepsilon)\} = \frac{1}{2\varepsilon}.$$

Let  $x^*$  be the optimal solution to the problem  $\max\{cx : x \in P(n+1, \varepsilon)\}$ . Without loss of generality, we can assume that the non-zero entries of  $x^*$  have indices in the set  $S \cup \{n+1\}$ , since all other indices have objective function coefficient equal to zero.

Let's first assume that  $x_{n+1}^* = 1$ . Then

$$\sum_{i \in S} c_i x_i^* = \sum_{i \in S} \frac{1}{2\varepsilon\sqrt{n}} x_i^* \leq \frac{1}{2\varepsilon} - 1,$$

where last inequality is implied by feasibility of  $x^*$ . It follows that  $cx^* \leq \frac{1}{2\varepsilon}$ .

If instead  $x_{n+1}^* = 0$ , we get  $\sum_{i \in S} c_i x_i^* \leq |S| \frac{1}{2\varepsilon\sqrt{n}} = \frac{1}{2\varepsilon}$ . On the other hand,

$$cx^S = (1 - \frac{\varepsilon}{2})\left(\frac{1}{2\varepsilon}\right) + \frac{4}{5} = \frac{1}{2\varepsilon} - \frac{1}{4} + \frac{4}{5} = \frac{1}{2\varepsilon} + \frac{11}{20}, \text{ hence}$$

$$(1 - \varepsilon) \frac{cx^S}{\max\{cx : x \in P(n+1, \varepsilon)\}} \geq (1 - \varepsilon)\left(1 + \frac{11}{10}\varepsilon\right) = 1 + \varepsilon\left(\frac{1}{10} - \frac{11}{10}\varepsilon\right) > 1$$

where last inequality follows because  $\frac{1}{10} - \frac{11}{10}\varepsilon > 0$ , since  $\varepsilon < \frac{1}{11}$ . □

**Claim 3.** For each  $S \neq S' \in \mathcal{S}$ , we have  $\frac{1}{2}x^S + \frac{1}{2}x^{S'} \in P(n+1, \varepsilon)$ . Hence,  $\mathcal{X}(n+1, \varepsilon)$  and  $P(n+1, \varepsilon)$  satisfy (iii).

*Proof.* In the following, we abbreviate  $P(n+1, \varepsilon)$  by  $P$ . Let  $x^{SS'} := \frac{1}{2}x^S + \frac{1}{2}x^{S'}$ . We show that  $x^{SS'} \in P$  by showing that  $x^{SS'} \leq \bar{x}$  (where the inequality is coordinate-wise) for a vector  $\bar{x} \in P$ , that is,  $x^{SS'}$  is *dominated* by a point  $\bar{x}$  inside the Knapsack polytope. Since the Knapsack polytope is down-monotone, the claim follows.

We now describe how to choose  $\bar{x}$ . Let  $\alpha := |S \cap S'| \leq \sqrt{n}/2 - 4$  and  $\beta := \lfloor (1 - 2\varepsilon)\sqrt{n} \rfloor$ . By our choice of  $n$  and  $\varepsilon$ , we have  $\beta - \alpha > 0$ . Consider the family  $\mathcal{K}$  all possible subsets of  $[n+1]$  of elements which:

- contain  $\{n+1\} \cup (S \cap S')$ ;
- contain  $\beta - \alpha$  elements from  $S \Delta S'$ .

Note that each of those sets has weight (with respect to our definition of  $P(n+1, \varepsilon)$ )

$$n + \frac{\beta}{2\varepsilon\sqrt{n}} \leq n + \frac{(1-2\varepsilon)\sqrt{n}}{2\varepsilon\sqrt{n}} \leq n + \frac{1}{2\varepsilon} - 1$$

hence it is a feasible Knapsack solution. Therefore, the incidence vector of each of those Knapsack solutions belongs to  $P$ , and we associate to each of them the multiplier  $\frac{4}{5} \cdot \frac{1}{|\mathcal{K}|}$ . Moreover, consider the set  $S \cup S'$ . This set also induces a feasible solution of  $P$ , since the total weight of the elements is at most  $1/\varepsilon < n$ , and therefore its incident vector is a point in  $P$ . We associate to this incident vector the multiplier  $1/5$ . By construction, those multipliers induce a convex combination of points from  $P$ , hence a point  $\bar{x} \in P$ .

We are left with showing that  $\bar{x}$  dominates  $x^{SS'}$ . Clearly, for any index  $i \notin S \cup S'$  we have  $\bar{x}_i = 0 = x_i^{SS'}$ . For the index  $n+1$  we have  $\bar{x}_{n+1} = x_{n+1}^{SS'} = 4/5$ . For indices  $i \in S \cap S'$  we have  $\bar{x}_i = 1 > (1 - \frac{\varepsilon}{2}) = x_i^{SS'}$ . Therefore, it only remains to analyze indices  $i \in S \Delta S'$ .

Observe that  $|S \Delta S'| = 2\sqrt{n} - 2\alpha$ , and hence that

$$\sum_{i \in S \Delta S'} \bar{x}_i = \frac{4}{5}(\beta - \alpha) + \frac{1}{5}(2\sqrt{n} - 2\alpha) \geq \frac{4}{5}((1 - 2\varepsilon)\sqrt{n} - 1 - \alpha) + \frac{1}{5}(2\sqrt{n} - 2\alpha).$$

Since  $\bar{x}_i = \bar{x}_j$  for  $i, j \in S \Delta S'$ , for each  $i \in S \Delta S'$ , one has:

$$\bar{x}_i \geq \frac{4}{5} \cdot \frac{(1 - 2\varepsilon)\sqrt{n} - 1 - \alpha}{2\sqrt{n} - 2\alpha} + \frac{1}{5}$$

and we want to show that the latter value exceeds  $x_i^{SS'}$ , that is, we would like the following inequality to hold

$$\frac{4}{5} \cdot \frac{(1 - 2\varepsilon)\sqrt{n} - 1 - \alpha}{2\sqrt{n} - 2\alpha} + \frac{1}{5} \geq \frac{1}{2} - \frac{\varepsilon}{4}$$

Now notice that

$$\begin{aligned} \frac{4}{5} \cdot \frac{(1 - 2\varepsilon)\sqrt{n} - 1 - \alpha}{2\sqrt{n} - 2\alpha} + \frac{1}{5} &\geq \frac{1}{2} - \frac{\varepsilon}{4} \\ &\Leftrightarrow \\ (1 - 2\varepsilon)\sqrt{n} - 1 - \alpha &\geq \left(\frac{3}{8} - \frac{5}{16}\varepsilon\right)(2\sqrt{n} - 2\alpha) \\ &\Leftrightarrow \\ \frac{\frac{1}{4} - \frac{11}{8}\varepsilon}{\frac{1}{4} + \frac{5}{8}\varepsilon}\sqrt{n} &\geq \alpha + \frac{1}{\frac{1}{4} + \frac{5}{8}\varepsilon} \end{aligned}$$

and the latter inequality is true since  $0 < \varepsilon < \frac{2}{27}$  implies that

$$\frac{\frac{1}{4} - \frac{11}{8}\varepsilon}{\frac{1}{4} + \frac{5}{8}\varepsilon} \geq \frac{1}{2} \quad \text{and} \quad 4 > \frac{1}{\frac{1}{4} + \frac{5}{8}\varepsilon},$$

therefore, recalling that  $\alpha \leq \sqrt{n}/2 - 4$ , we get

$$\frac{\frac{1}{4} - \frac{11}{8}\varepsilon}{\frac{1}{4} + \frac{5}{8}\varepsilon} \sqrt{n} \geq \frac{1}{2} \sqrt{n} \geq \alpha + 4 \geq \alpha + \frac{1}{\frac{1}{4} + \frac{5}{8}\varepsilon}$$

as claimed.  $\square$

### 3 An upper bound for down-monotone polytopes

In this section we prove Theorem 2.

Let  $P \subseteq \mathbb{R}_+^n$  be a down-monotone polytope. By scaling it, we can suppose that  $P \subseteq [0, 1]^n$ . For any  $\varepsilon$  with  $1/2 \geq \varepsilon > 0$ , set  $\gamma := \varepsilon/(1 - \varepsilon) \leq 1$ . We will approximate  $P$  with inequalities whose coefficients belong to the following set:

$$S := \{0\} \cup \{ \lfloor (1 + \frac{\gamma}{2})^\ell \rfloor : \ell = 0, 1, \dots, \lfloor \log_{(1+\frac{\gamma}{2})} \left( \frac{5n}{\gamma} \right) \rfloor \}.$$

Note that  $|S| \leq 1 + \min\{ \lfloor \log_{(1+\frac{\gamma}{2})} \left( \frac{5n}{\gamma} \right) \rfloor, \frac{5n}{\gamma} \}$ , and we have

$$\log_{(1+\frac{\gamma}{2})} \left( \frac{5n}{\gamma} \right) = O \left( \frac{\log(5n/\varepsilon)}{\log(1 + \frac{\varepsilon}{2})} \right) = O \left( \frac{\log n/\varepsilon}{\varepsilon} \right),$$

where the last equality follows from the fact that  $\frac{\log(1+\frac{\varepsilon}{2})}{\varepsilon} = \Theta(1)$  for  $\varepsilon \rightarrow 0$ , which can be verified using e.g. l'Hôpital's rule. We conclude  $|S| = O(\min\{ \frac{\log(n/\varepsilon)}{\varepsilon}, \frac{n}{\varepsilon} \})$ , and in particular  $|S| = O(\log n)$  for any constant  $\varepsilon$ .

Let now  $Q$  be the polytope obtained by intersecting  $1 \geq x \geq 0$  with all the inequalities of the form  $\tilde{c}x \leq \max\{\tilde{c}y : y \in P\}$  with  $\tilde{c}_i \in S$  for  $i \in [n]$ . That is:

$$Q := \{x \in \mathbb{R}^n : 0 \leq x \leq 1, \tilde{c}x \leq \max\{\tilde{c}y : y \in P\} \text{ for all } \tilde{c} \in \mathbb{R}^n \text{ with } \tilde{c}_i \in S \forall i \in [n]\}.$$

In order to prove Theorem 2, we are left to show the following.

**Lemma 1.**  *$Q$  is an  $\varepsilon$ -approximated formulation for  $P$ .*

*Proof.* Clearly  $P \subseteq Q$ . Consider any vector  $c \in \mathbb{R}^n$ . We have to show that

$$\beta := \max\{cx : x \in P\} \geq (1 - \varepsilon) \max\{cx : x \in Q\}.$$

As  $P$  is down-monotone, we can assume  $c \geq 0$ . Define  $\tilde{c}$  as follows. Let  $K = \|c\|_\infty \frac{\gamma}{5n}$  and  $G := \{i \in [n] : c_i < K\}$ . Set  $\tilde{c}_i = 0$  for  $i \in G$  and  $\tilde{c}_i = \lfloor (1 + \frac{\gamma}{2})^{\lfloor \log_{(1+\frac{\gamma}{2})} (\frac{c_i}{K}) \rfloor} \rfloor$  otherwise. Note that we have  $\tilde{c}_i \in S$  for all  $i \in [n]$ .

Let  $\tilde{\beta} := \frac{\beta}{K}$ . We now show

$$\max\{cx : x \in Q\} \leq \max\{cx : \tilde{c}x \leq \tilde{\beta}, 0 \leq x \leq 1\} \leq (1 + \gamma)\beta = \beta/(1 - \varepsilon),$$

which immediately concludes the proof. The equality follows by definition of  $\gamma$ . The first inequality follows easily by observing that the constraint  $\tilde{c}x \leq \tilde{\beta}$  is a valid constraint for  $Q$ , as  $\tilde{c}_i \in S$  for all  $i \in [n]$  and for  $x \in P$  one has

$$\sum_i \tilde{c}_i x_i = \sum_{i \notin G} \tilde{c}_i x_i \leq \sum_{i \notin G} (1 + \frac{\gamma}{2})^{\log_{(1+\frac{\gamma}{2})} (\frac{c_i}{K})} x_i \leq \frac{1}{K} \sum_i c_i x_i \leq \frac{\beta}{K} = \tilde{\beta}.$$

For the second inequality, let  $x^* \in [0, 1]^n$  be such that  $cx^* > (1 + \gamma)\beta$ . We show that  $\tilde{c}x^* > \tilde{\beta}$ , concluding the proof. We have:

$$\begin{aligned}\tilde{c}x^* &= \sum_{i \notin G} \tilde{c}_i x_i^* \\ &\geq \sum_{i \notin G} \frac{2}{2 + \gamma} \frac{c_i}{K} x_i^* - n\end{aligned}\tag{2}$$

$$\begin{aligned}&= \frac{2}{2 + \gamma} \frac{1}{K} \left( \sum_{i \in [n]} c_i x_i^* - \sum_{i \in G} c_i x_i^* \right) - n \\ &\geq \frac{2}{2 + \gamma} \frac{1}{K} cx^* - \frac{2}{2 + \gamma} \frac{1}{K} Kn - n\end{aligned}\tag{3}$$

$$\begin{aligned}&> \frac{2}{2 + \gamma} \frac{1}{K} \beta(1 + \gamma) - \frac{2n}{2 + \gamma} - n \\ &= \tilde{\beta} + \frac{\gamma}{2 + \gamma} \tilde{\beta} - \frac{2n}{2 + \gamma} - n \\ &\geq \tilde{\beta},\end{aligned}\tag{4}$$

where (2) is implied by  $x^* \in [0, 1]^n$  and, for  $i \notin G$ ,

$$\tilde{c}_i \geq \left(1 + \frac{\gamma}{2}\right)^{\log_{(1+\frac{\gamma}{2})}(\frac{c_i}{K})-1} - 1 = \frac{c_i}{K} \left(1 + \frac{\gamma}{2}\right)^{-1} - 1 = \frac{2}{2 + \gamma} \frac{c_i}{K} - 1;$$

(3) follows from the fact that  $c_i < K$  for  $i \in G$ ,  $|G| \leq n$ , and  $x^* \leq 1$ ; and (4) holds because

$$\frac{\gamma}{2 + \gamma} \tilde{\beta} - \frac{2n}{2 + \gamma} - n \geq 0 \Leftrightarrow \tilde{\beta} \geq \frac{4n + \gamma n}{\gamma} \Leftrightarrow \beta \geq Kn \frac{(4 + \gamma)}{\gamma} = \|c\|_{\infty} \frac{4 + \gamma}{5},$$

and the latter is true because our choices of parameters imply  $\beta \geq \|c\|_{\infty}$  and  $\gamma \leq 1$ .  $\square$

**Acknowledgments.** We thank Carsten Moldenhauer for useful discussions.

## References

- [1] E. Balas. Disjunctive programming. *Annals of Discrete Mathematics* 5: 3–51, 1979.
- [2] D. Bienstock. Approximate formulations for 0–1 knapsack sets. *Oper. Res. Lett.* 36: 317–320, 2008.
- [3] D. Bienstock and B. McClosky. Tightening simple mixed-integer sets with guaranteed bounds. *Math. Program., Ser. A* 133: 337–363, 2012.
- [4] G. Braun and S. Pokutta. The matching polytope does not admit fully-polynomial size relaxation schemes. *Proceedings of 26th SODA*: 837–846, 2015.
- [5] T.H. Cormen, C.E. Leiserson, R.L. Rivest, and C. Stein. Introduction to Algorithms, MIT press, 2010.
- [6] O.H. Ibarra and C.E. Kim. Fast Approximation Algorithms for the Knapsack and Sum of Subsets Problems. *J. ACM* 22: 463–468, 1975.
- [7] V. Kaibel and S. Weltge. Lower Bounds on the Sizes of Integer Programs without Additional Variables. *Mathematical Programming, series B*. doi:10.1007/s10107-014-0855-0, 2014.

- [8] R. M. Karp. Reducibility Among Combinatorial Problems. In R. E. Miller and J. W. Thatcher (editors), *Complexity of Computer Computations*: 85–103, 1972.
- [9] H. Kellerer, U. Pferschy, and D. Pisinger. Knapsack problems. Springer, 2004.
- [10] S. G. Kolliopoulos and Y. Moysoglou. Exponential lower bounds on the size of approximate formulations in the natural encoding for Capacitated Facility Location. Manuscript, 2013. Available online at <http://arxiv.org/abs/1312.1819>.
- [11] E.L. Lawler. Fast Approximation Schemes for Knapsack Problems. *Proc. 18th FOCS*: 206–213, 1977.
- [12] G. Nemhauser, L. Wolsey. Integer and combinatorial optimization. Wiley interscience series in discrete mathematics and optimization, 1988.
- [13] N. Nisan and A. Wigderson. Hardness vs randomness. *J. Comput. Syst. Sci.* 49 (2): 149–167, 1994.
- [14] M. Van Vyve. Polynomial relaxation schemes for knapsack. Manuscript, 2005.
- [15] M. Van Vyve and L. Wolsey. Approximate extended formulations. *Math. Program., Ser. B* 105: 501–522, 2006.